THE LECTURE 17

The Gram-Schmidt Algorithm

In the last lecture I showed how one could break a vector \( \mathbf{v} \) up into two orthogonal components; with one component lying in a given subspace \( W \) and another component lying in the subspace \( W^\perp \) that is orthogonal to \( W \). The procedure was to

- choose a basis \( B_W = \{ \mathbf{b}_1, \ldots, \mathbf{b}_k \} \) for \( W \)
- find a basis \( B_{W^\perp} = \{ \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n \} \) for \( W^\perp \)
- combine \( B_W \) with \( B_{W^\perp} \) to form a basis \( B = \{ \mathbf{b}_1, \ldots, \mathbf{b}_n \} \) for \( \mathbb{R}^n \)
- find the coordinate vector \( \mathbf{v}_B \) of \( \mathbf{v} \) with respect to \( B \) and then throw away the components along the vectors \( \{ \mathbf{b}_{k+1}, \ldots, \mathbf{b}_n \} \)

Today we develop a more systematic approach that

**Theorem 17.1.** Let \( \{ \mathbf{v}_1, \ldots, \mathbf{v}_k \} \) be a set of mutually orthogonal non-zero vectors. Then the vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_k \) are linearly independent.

**Proof.**

Suppose

\[
  c_1 \mathbf{v}_1 + \cdots + c_k \mathbf{v}_k = \mathbf{0}
\]

Then for each \( i = 1, \ldots, k \) we have

\[
  0 = \mathbf{0} \cdot \mathbf{v}_i = c_1 \mathbf{v}_1 \cdot \mathbf{v}_i + c_2 \mathbf{v}_2 \cdot \mathbf{v}_i + \cdots + c_k \mathbf{v}_k \cdot \mathbf{v}_i = c_i \| \mathbf{v}_i \|^2 \implies c_i = 0
\]

So we cannot satisfy (1) without each \( c_i = 0 \). Hence the vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_k \) are linearly independent.

**Corollary 17.2.** Any set of \( n \) mutually orthogonal non-zero vectors will be a basis for \( \mathbb{R}^n \).

Now suppose \( B = \{ \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k \} \) is a basis for some subspace \( W \) of \( \mathbb{R}^n \). From this basis we can systematically construct an orthogonal basis for \( W \); that is a basis for which all the vectors are orthogonal.

Before we get started, let’s recall that given any vectors \( \mathbf{a} \) and \( \mathbf{v} \) we have a decomposition of \( \mathbf{v} \)

\[
\mathbf{v} = \mathbf{v}_a + \mathbf{v}_{a^\perp}
\]

where \( \mathbf{v}_a \) is the component of \( \mathbf{v} \) along the direction of \( \mathbf{a} \) and \( \mathbf{v}_{a^\perp} \) is the component of \( \mathbf{v} \) along a direction perpendicular to \( \mathbf{v} \). Moreover, we have the following formula for \( \mathbf{v}_a \)

\[
\mathbf{v}_a = \frac{\mathbf{a} \cdot \mathbf{v}}{\mathbf{a} \cdot \mathbf{a}}
\]

Combining (2) and (3) we have a formula for \( \mathbf{v}_{a^\perp} \) as well

\[
\mathbf{v}_{a^\perp} = \mathbf{v} - \frac{\mathbf{a} \cdot \mathbf{v}}{\mathbf{a} \cdot \mathbf{a}}
\]
Remark 17.3. Note that if \( \mathbf{a} \) and \( \mathbf{v} \) are linearly independent then \( \mathbf{v}_{\perp} \neq \mathbf{0} \) : because it is a linear combination of two linearly independent vectors with at least one coefficient, the coefficient of \( \mathbf{v} \), non-zero. Note also that from (3)

\[
\mathbf{a} \cdot \mathbf{v}_{\perp} = \mathbf{a} \cdot \mathbf{v} + \frac{\mathbf{a} \cdot \mathbf{v}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a} = 0
\]
as expected.

Okay, here’s how we generate an orthogonal basis. Set

\[
\mathbf{o}_1 = \mathbf{b}_1
\]
and then

\[
\mathbf{o}_2 = \mathbf{b}_2 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_2}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1
\]

By construction, \( \mathbf{o}_1 \) and \( \mathbf{o}_2 \) are perpendicular, non-zero and linearly independent. Now let

\[
\mathbf{o}_3 = \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2
\]
The vector \( \mathbf{o}_3 \) is non-zero because it is a linear combination of the basis vectors \( \mathbf{b}_1, \mathbf{b}_2 \) and \( \mathbf{b}_3 \) with at least one non-zero coefficient. Moreover

\[
\begin{align*}
\mathbf{o}_1 \cdot \mathbf{o}_3 &= \mathbf{o}_1 \cdot \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2 = \mathbf{o}_1 \cdot \mathbf{b}_3 - \mathbf{o}_1 \cdot \mathbf{b}_3 = 0 \\
\mathbf{o}_2 \cdot \mathbf{o}_3 &= \mathbf{o}_2 \cdot \mathbf{b}_3 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_3}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_3}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2 = \mathbf{o}_2 \cdot \mathbf{b}_3 - \mathbf{o}_2 \cdot \mathbf{b}_3 = 0
\end{align*}
\]
and so \( \{ \mathbf{o}_1, \mathbf{o}_2, \mathbf{o}_3 \} \) are mutually perpendicular non-zero vectors, and so linearly independent.

We can continue in this fashion to construct more and more linearly independent orthogonal vectors. For example, \( \mathbf{o}_4 = \mathbf{b}_4 - \frac{\mathbf{o}_1 \cdot \mathbf{b}_4}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_4}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2 - \frac{\mathbf{o}_3 \cdot \mathbf{b}_4}{\mathbf{o}_3 \cdot \mathbf{o}_3} \mathbf{o}_3 \)

In the end, when we reach \( \mathbf{b}_k \) this process terminates with

\[
\mathbf{o}_k = \mathbf{b}_k - \frac{\mathbf{o}_1 \cdot \mathbf{b}_k}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 - \frac{\mathbf{o}_2 \cdot \mathbf{b}_k}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2 - \cdots - \frac{\mathbf{o}_{k-1} \cdot \mathbf{b}_k}{\mathbf{o}_{k-1} \cdot \mathbf{o}_{k-1}} \mathbf{o}_{k-1}
\]
and we arrive at a set of \( k \) linearly independent, mutually orthogonal vectors \( \{ \mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_k \} \)

The basis \( \{ \mathbf{o}_1, \ldots, \mathbf{o}_k \} \) obtained by the above algorithm, however, is not an orthonormal basis. That is to say, although mutually orthogonal by construction, the vectors \( \mathbf{o}_i \) do not necessarily have the length 1. In fact, it’s rather unlikely that \( ||\mathbf{o}_i|| = 1 \). But there is an easy fix for this. All we have to do is divide each of the orthogonal basis vectors \( \mathbf{o}_i \) by their lengths \( ||\mathbf{o}_i|| = \sqrt{\mathbf{o}_i \cdot \mathbf{o}_i} \) to get a set of \( k \), mutually orthogonal, linearly independent vectors, all of length 1:

\[
\begin{align*}
\mathbf{o}_1 &\rightarrow \mathbf{n}_1 = \frac{1}{\mathbf{o}_1 \cdot \mathbf{o}_1} \mathbf{o}_1 \\
\mathbf{o}_2 &\rightarrow \mathbf{n}_2 = \frac{1}{\mathbf{o}_2 \cdot \mathbf{o}_2} \mathbf{o}_2 \\
&\quad \vdots \\
\mathbf{o}_k &\rightarrow \mathbf{n}_k = \frac{1}{\mathbf{o}_k \cdot \mathbf{o}_k} \mathbf{o}_k
\end{align*}
\]

Example 17.4. Find an orthonormal basis for the subspace

\[
W = \text{span}\left( \left[ 1, -1, 1, 0, 0 \right], \left[ -1, 0, 0, 0, 1 \right], \left[ 0, 0, 1, 0, 1 \right] \right)
\]
of \( \mathbb{R}^5 \).
First we need a basis for $W$.

$$\begin{bmatrix}
1 & -1 & 1 & 0 & 0 \\
-1 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 \\
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & -1 & 1 & 0 & 0 \\
0 & -1 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 \\
\end{bmatrix}$$

This last matrix is in row echelon form with no non-zero rows. From this short calculation we see that the original three vectors are linearly independent and so will constitute a basis for $W$. We can thus use $B = \{b_1, b_2, b_3\}$ with

- $b_1 = [1, -1, 1, 0, 0]$  
- $b_2 = [-1, 0, 0, 0, 1]$  
- $b_3 = [0, 0, 1, 0, 1]$  

as an initial basis to start the Gram-Schmidt orthogonalization process.

Thus, we set

$$o_1 = b_1 = [1, -1, 1, 0, 0]$$

$$\Rightarrow \|o_1\|^2 = 3$$

$$\Rightarrow n_1 = \left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0, 0\right]$$

Next we compute $o_2$,

$$o_2 = b_2 - \frac{o_1 \cdot b_2}{o_1 \cdot o_1} o_1$$

$$= [-1, 0, 0, 0, 1] - \frac{-1}{3} [1, -1, 1, 0, 0]$$

$$= \left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 0, 1\right]$$

We have

$$\|o_2\|^2 = \frac{4}{9} + \frac{1}{9} + \frac{1}{9} + 1 = \frac{5}{3}$$

$$\Rightarrow n_2 = \sqrt{\frac{3}{5}} \left[-\frac{2}{3 \sqrt{3}}, -\frac{1}{3 \sqrt{3}}, \frac{1}{3 \sqrt{3}}, 0, 1\right]$$

Finally,

$$o_3 = b_3 - \frac{o_1 \cdot b_3}{o_1 \cdot o_1} o_1 - \frac{o_2 \cdot b_2}{o_2 \cdot o_2} o_2$$

$$= [0, 0, 1, 0, 1] - \frac{1}{3} [1, -1, 1, 0, 0] - \frac{1}{2} [-1, 0, 0, 0, 1]$$

$$= \left[\frac{1}{6}, \frac{1}{3}, \frac{2}{3}, 0, \frac{1}{2}\right]$$

and

$$\|o_3\|^2 = \frac{1}{36} + \frac{1}{9} + \frac{1}{4} + \frac{1}{4} = \frac{5}{6}$$

so

$$n_3 = \sqrt{\frac{6}{5}} \left[\frac{1}{6}, \frac{1}{3}, \frac{2}{3}, 0, \frac{1}{2}\right]$$

Thus,

$$B' = \left\{ \left[\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, 0, 0\right], \sqrt{\frac{3}{5}} \left[-\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, 0, 1\right], \sqrt{\frac{6}{5}} \left[\frac{1}{6}, \frac{1}{3}, \frac{2}{3}, 0, \frac{1}{2}\right] \right\}$$

will be an orthonormal basis for $W$. 